(12)

ANALYSIS, MODELING, AND SIGNAL DETECTION FOR A SET OF PASSIVE SONAR DATA*

C.R. Baker, H. Cherifi, T.G. Dankel, and M.R. Frey

Department of Statistics
University of North Carolina
Chapel Hill, NC 27599

LISS 50

August, 1992 and February, 1988

*Work supported by ONR Contract N00014-86-K-0039.

Table of Contents

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for Public Release: Distribution Unlimited
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Department of Statistics 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION
6c. ADDRESS (City, State and ZIP Code) University of North Carolina Chapel Hill, North Carolina 27514	7b. ADDRESS (City, State and ZIP Code)
8a. NAME OF FUNDING/SPONSORING ORGANIZATION Office of Naval Research 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER N00014-86-K-0039 and N00014-89-J-1175
8c. ADDRESS (City, State and ZIP Code) Statistics & Probability Program Arlington, VA 22217	10. SOURCE OF FUNDING NOS. PROGRAM ELEMENT NO. / PROJECT NO. / TASK NO. / WORK UNIT NO.
11. TITLE (Include Security Classification) Analysis, Modeling, and Signal (cont.)	
12. PERSONAL AUTHOR(S) C.R. Baker, H. Cherifi, T.G. Dankel and M.R. Frey	
13a. TYPE OF REPORT TECHNICAL 13b. TIME COVERED FROM _____ TO _____	14. DATE OF REPORT (Yr., Mo., Day) Feb. 1988 & Aug. 1992 15. PAGE COUNT 34
16. SUPPLEMENTARY NOTATION	

17. COSATI CODES		
FIELD	GROUP	SUB. GR.

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
Passive sonar, Sonar signal detection,
Sonar data analysis

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

A computational evaluation of a set of single-hydrophone passive sonar data is summarized. Included are results on signal detection, statistical characterization (especially normality vs. non-normality), and fits to a mixture-of-normals distribution.

TITLE CONT.: Detection for a Set of Passive Sonar Data

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☒ DTIC USERS ☐	21. ABSTRACT SECURITY CLASSIFICATION
22a. NAME OF RESPONSIBLE INDIVIDUAL C.R. Baker	22b. TELEPHONE NUMBER (Include Area Code) (919) 962-2189 22c. OFFICE SYMBOL

DD FORM 1473, 83 APR EDITION OF 1 JAN 73 IS OBSOLETE.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

Note to the Reader

(by C. R. Baker, August, 1992)

This report has previously been provided to Drs. C.E. Persons and D.H. Gingras of NOSC (Naval Ocean Systems Center), in 1988. Recent interest by other personnel at NOSC (now NCCOSC) having access to the report has indicated that it may be of wider interest. We are thus publishing it as a LISS report, essentially unchanged except for omission of possibly-sensitive frequency values and a lofargram, and with some explanatory notes added on figures that were originally in color. The analysis is on data taken from a single hydrophone in the deep ocean.

As regards the signal detection results contained in the report, they have been extended by work reported in reference {1} below. The later work used the NOSC data described here to evaluate new detection algorithms derived from results in {2}. That work, and approximations described in {3}-{5}, led to the representation of signal-plus-noise as a filtered diffusion when the noise is Gaussian, apparently a new concept in representation of sonar data, and one which has led to the promising algorithms described in {1} and {3}-{5}.

A second aspect of this report is on statistical analysis of the NOSC data. This led to characterizing the data as nonGaussian signal-plus-noise and Gaussian noise. Some of our more recent work with this same data set has indicated that the data properties are sensitive to the particular subset being tested; for example, some subsets of the noise fail to pass tests for normality.

As will be seen, this report strongly recommended that the Navy initiate a serious study on detection of nonGaussian signals in Gaussian noise. Subsequently, Dr. Persons initiated an SBIR task statement on this topic, which was approved and included in the D0D SBIR solicitation of October, 1988. A Phase I SBIR contract was let in 1989 for a survey of existing approaches, and this led to a Phase II contract for algorithm evaluation, similar to the work reported in {1} but with extensions to arrays. The new results on this problem described in {3}, and discussed far more completely in {4}, have apparently been prominently employed by the contractor for the SBIR work. In fact, the PI for the contractor awarded the SBIR contracts stated (in the course of preparing his proposal for the Phase II award) that he considered the approach developed in {4} as being the most promising for this problem.

A third aspect of this report is its venture into classification. This is by modeling the data as a Gaussian mixture (spherically-invariant) process: $Y = AX$, where X is a Gaussian process and A is a positive random variable independent of X. The parameters describing the probability distribution of the "mixing random variable" A are the number K of values that A can assume, and $\{(a_i, p_i): 1 \le i \le K\}$, where $A = a_i$ with probability p_i. Although we investigated the method of moments as a procedure for estimating these parameters, most of our work (and all of that reported here) utilized a modification of the EM (Expectation-Maximization) algorithm {6}. In the literature, however, the number K of values that the mixing random variable can assume is typically assumed known. Since this is a very important parameter, and since the EM algorithm is known to be sensitive to its initial conditions, we devised a data-based procedure for estimating the initial distribution of A, including the value of K. Procedures for estimating K were also incorporated into the algorithm.

The work reported here consisted of estimating the parameters of A for data that was not accepted as normal by statistical tests for normality. As could be expected from physical reasoning, we found that the number of values of A with signal present was significantly larger than with signal absent, for data in the low-frequency region. Differences in the (a_i, p_i) values were also observed. It seems reasonable that the differences exhibited would be greater if the characterization were applied to an estimate of the signal, rather than to the combined signal-plus-noise process. The same behavior seems likely in discriminating by active sonar between a complex distributed target (e.g., submarine) and a less-complex target (e.g., rocks, marine organisms) having a similar reflected energy level.

A natural approach is to apply this characterization to signal estimates obtained from the procedures described in {1} and {3}-{5}, an area of current interest to us.

We now turn to a very brief summary of some of our other work related to this report, and in particular to the use of Gaussian mixtures. In {3}, the detection of a known signal imbedded in such noise is considered. The log-likelihood ratio is given, two constant-false-alarm-probability detection algorithms are obtained, and the effect of sampling rate is analyzed. That paper also demonstrates that the Class A impulsive noise model of David Middleton is a special case of a Gaussian mixture.

In {7}, the likelihood ratio for detection of a random signal in Gaussian mixture noise is obtained for the continuous-time problem, along with the capacity of a communication channel perturbed by such noise. The paper {7} is a purely mathematical treatment; discrete-time algorithms for the detection of a random signal in Gaussian mixture noise are derived in {4}, currently the only source for those algorithms. The publication date of the book to contain {4}, which was given as 1990 when the paper was solicited, is at present not known to us. However, copies of the paper can be obtained from the Department of Statistics at UNC-Chapel Hill.

The thesis {8} applied the EM algorithm to model simulated data as a Gaussian mixture. Other work, not published, has included the modeling of under-ice sonar data furnished by NUWC (Dr. Roger Dwyer) as a Gaussian mixture. An example of the results obtained is the noise distribution used to obtain the performance curves shown in Figure 1 of {3}.

The results obtained in {2}, which lead to the representation of the signal-plus-noise process as a filtered diffusion when the noise is Gaussian, and thereby to the new algorithms described in {3}-{5}, provide a promising approach for detection of a random signal in Gaussian and a large class of nonGaussian noise, particularly for signals that are broadband and/or nonstationary and whose statistical properties are partially or wholly unknown. At the same time, physical rationale furnishes an incentive to continue work with Gaussian-mixture noise processes. The results of {7} and {4} provide a means of combining these two areas in the development of new algorithms for applications where the noise is either a Gaussian mixture or contains an additive component of this type. Moreover, the use of such a model appears to have potential as an aid to classification, and results obtained in the detection problem should be useful in the extension to classification.

Acknowledgements

The computer programs for the statistical tests, detection algorithms, algorithm evaluation methods, and modeling algorithms used to obtain the results in this report were written by LISS personnel on a computer system that was new to all of us. This was carried out without benefit of any software package other than basic programming languages, without a systems administrator, and without a systems programmer. Over 10,000 lines of programs were written during the 1984-1987 period alone. This major effort required sustained contributions from a number of people. The report was prepared by C. R. Baker and M. R. Frey, but all of the listed authors made contributions that were essential to the work, and important contributions were also made by G. Demuth, H. Hucke, J. Strobel, and L. Trimble.

References

1. C.R. Baker, M.R. Frey, and C.E. Persons, Computational evaluations of new optimum detection algorithms, *USN J. Underwater Acoustics, 41,* 457-476 (1991).

2. _____ and A.F. Gualtierotti, Discrimination with respect to a Gaussian process, *Probability Theory and Related Fields, 71,* 159-182 (1986).

3. _____, Likelihood ratios and signal detection for nonGaussian processes, *Stochastic Processes in Underwater Acoustics,* Lecture Notes in Control and Information Science, Springer-Verlag, *85,* 154-180, Berlin (1986).

4. _____, Likelihood-ratio detection of stochastic signals, *Advances in Statistical Signal Processing. Volume 2. Signal Detection,* H.V. Poor and J.B. Thomas, eds., JAI Press (to appear). Available from Dept. of Statistics, University of North Carolina, Chapel Hill, N.C. 27599-3260, as tech. rept. LISS 40 (October, 1989).

5. _____, Algorithms for detection of stochastic signals imbedded in Gaussian noise, LISS 48 (March, 1991), Dept. of Statistics, University of North Carolina, Chapel Hill, N.C. 27599-3260.

6. D.M. Titterington, A.F.M. Smith, and U.E. Makov, *Statistical Analysis of Finite Mixture Distributions,* Wiley (1950).

7. C. R. Baker and A. F. Gualtierotti, Absolute continuity and mutual information for Gaussian mixtures, *Stochastics and Stochastics Reports, 39,* 139-157 (1992).

8. H. P. Hucke, Investigation of an iterative maximum likelihood method for estimating mixture distributions with zero mean, Thesis, Department of Statistics, Universtiy of North Carolina at Chapel Hill (June, 1985).

0. GENERAL DISCUSSION

A. Introduction

This report summarizes results of a computational study of underwater acoustics data furnished by the Naval Ocean Systems Center (Dr. C.E. Persons and Dr. D.H. Gingras). The data contains several minutes of apparent noise-only data, followed by several minutes of apparent signal-plus-noise, and then several more minutes of apparent noise-only data. The source of the signal data was not furnished.

The purpose of the study was two-fold: (1) to study the statistical properties of the data; (2) to evaluate performance of several detection algorithms. To summarize, the noise data was found to be strongly Gaussian, the signal-plus-noise data nonGaussian. This points out a clear need for the development of detection algorithms for such data.

The investigation in each area will now be briefly summarized.

B. Summary of Statistical Analysis

One objective of the statistical analysis was to determine if the non-Gaussian data samples could be modeled as a spherically-invariant random variable (SIRV): a random variable of the form $X = AY$, where Y is a Gaussian random variable and A is a "mixing" random variable independent of Y. We also refer to a SIRV as a mixture-of-normals. Such a random variable has a density function that appears normal in its general form, but which is actually more peaked and with heavier tails than a true Gaussian random variable with the same mean and variance (see Appendix 2). A discussion of such a model is contained in [1]. One well-known model of Gaussian-plus-impulsive noise that is actually a SIRV is the "Class A" model described in [2].

In the study summarized here, all data samples were subjected to four tests of normality. Samples which failed at least two of these tests were modeled as having a spherically-invariant distribution. Using a modified maximum-likelihood algorithm, the distribution of the r.v. A (in the representation $X = AY$) was estimated. The most interesting results of this work can be summarized as follows:

(1) In the low-frequency band and mid-frequency the noise-only data appeared to be Gaussian, with 6 out of 11 sample vectors in each segment passing all tests for normality. An even stronger indication of normality was the estimated number of components for the mixing random variable A. For Gaussian data, this should be equal to unity. This was the value estimated for 10 of 11 sample vectors in the low-frequency band, and for 8 of 11 in the mid-frequency band.

In these same frequency bands, the signal-plus-noise data did not appear to be Gaussian, with 9 out of 11 sample vectors in each segment failing 2 of the 4 normality tests.

In the high-frequency band the situation was not as clear. Noise-only data appeared to be tending toward nonGaussian, while the signal-plus-noise data appeared to be tending toward Gaussian.

One may view the number of estimated components (possible values) of the random variable A as an indicator of both the degree of non-normality and the complexity of the data. In the low-frequency band, the estimated number of components for the noise-only data was one (i.e., Gaussian data) for 10 of the 11 data samples. By contrast, the signal-plus-noise data had a median of 3 for the estimated number of components for A.

Focusing on the low-frequency band, the noise-only data clearly appeared to be Gaussian. The signal-plus-noise seemed just as clearly to be nonGaussian. Thus, the signal detection problem for this data can be viewed as detection of a nonGaussian signal imbedded in Gaussian noise.

C. Summary of Signal Detection Study

The signal detection study involved the empirical evaluation of four signal detection algorithms. Two were quadratic detectors: the optimum detector under the assumption that both the noise and signal-plus-noise were Gaussian, and the optimum deflection criterion detector [3]. The other two were detectors that assume that the signal can be modeled as a deterministic function, leading to a linear operation on the data. One of these detectors is the matched filter; the other is the optimum detector assuming that the noise is a spherically-invariant random process: $\underline{X} = A\underline{Y}$, where $\underline{X}$ and $\underline{Y}$ are random vectors, $\underline{Y}$ is Gaussian, and A is a random variable independent of $\underline{Y}$. For the two linear detectors, the returned signal waveform was estimated by using the sample mean of the signal-plus-noise ensemble.

The results of these evaluations show that the two quadratic detectors were considerably superior to the two linear detectors, and that the performance of the two quadratic detectors was about equal.

D. Suggestions for Further Investigations

1. Statistical Analysis

This single tape may represent too small a data sample to permit general conclusions to be drawn for the type of ship and environment producing the data. However, in order to carry out a more complete statistical analysis of such data, several extensions to the work reported here are desirable.

First, algorithms are needed for testing whether or not the univariate data, when nonGaussian, can be accepted as spherically-invariant (mixture of normals). This seems to be a reasonable hypothesis, based on the data observed here, but a quantitative test of this hypothesis should be implemented and applied.

A second major extension would be to improve the flexibility of the algorithm used to estimate the distribution of the mixing random variable A in the representation $X = AY$ previously discussed. The current algorithm is such that the estimated number of components between iterations cannot increase; this restriction needs to be removed. Without this restriction, it is likely that the estimated number of components for the mixing random variable A will increase, in the signal-plus-noise situation. There is likely to be little change in the number of components for the noise-only data, as it seems rather clearly to be Gaussian.

The third major extension that should be considered is to extend the nonGaussian modeling effort from one of modeling the univariate data to one of modeling multivariate data. Two possibilities are the spherically-invariant random process ($\underline{X} = A\underline{Y}$, $\underline{Y}$ a Gaussian vector, A a random variable independent of $\underline{Y}$) or a generalized spherically-invariant random process, or GSIRP ($\underline{X} = A\underline{Y}$, Y a Gaussian vector, A a diagonal matrix whose non-zero components are random variables independent of $\underline{Y}$).

Finally, an interesting possibility is to model the data as a filtered diffusion: $Z_n = \sum_{j=0}^{n-1} A(n,Z_j) + W_n$, with $X_n = \sum_{j=0}^{n-1} F(n,j)\Delta Z_j$, $\underline{X}$ the data, $\underline{W}$ white Gaussian noise, A a drift function, F a linear filter, $\Delta Z_n = Z_j - Z_{j-1}$. Such a model has been considered in [1] and [4].

2. Signal Detection

As noted, the signal detection problem for this data appears to be that of detecting a broad-band nonGaussian signal imbedded in Gaussian noise. The work reported here did not include the performance evaluation of an optimum (likelihood ratio) detector for this nonGaussian detection problem.

Of course, in order to obtain a likelihood ratio detector, one must know the distributions of the signal-plus-noise process. Three reasonable possibilities for the signal-plus-noise process that merit investigation are the following:

a) $\underline{X} = A\underline{S} + \underline{N}$, where $\underline{N}$ is Gaussian, $\underline{S}$ is Gaussian, A is a random variable independent of both the random vectors $\underline{N}$ and $\underline{S}$;

b) $\underline{X} = A\underline{S} + \underline{N}$, where $\underline{S}$ and $\underline{N}$ are as in a), and A is a diagonal matrix whose non-zero components are random variables, each independent of $\underline{S}$;

c) $\underline{X}$ is a filtered diffusion, as discussed in D.1:

$$X(n) = \sum_{j=0}^{n-1} F(n,j)\Delta Z_j,$$

$$Z(n) = \sum_{j=0}^{n-1} A(n,Z_j) + W_j,$$

with $\underline{W}$ white Gaussian noise, A a drift function, F a linear filter. A detection algorithm for this model has been obtained [1]. Its implementation requires knowledge of the matrix F and drift function A. The major difficulty is estimation of A, and some work on this problem has been done by the ISS research program.

The computational results summarized here suggest the investigation of a detection algorithm based on estimation of the number of components of the random variable A in the model X = AY discussed in B. above. As discussed in Section B above, the median number of components for the signal-plus-noise data was 3 in the low-frequency band, compared to 1 for the noise-only data. As noted in D.1, the estimated number of components for A is likely to increase (for the signal-plus-noise data) once the computer program is modified as described.

Of course, it may be that none of these algorithms will eventually provide a likelihood ratio detector for the type of detection problem represented by the NOSC Tape data. However, if this data set is representative, the problem is clearly (in the low-frequency and middle-frequency bands of the data) one of discriminating between Gaussian noise and nonGaussian signal-plus-noise. The computational results on modeling also indicate that the univariate data exhibits significant statistical differences between the noise data and the signal-plus-noise data (e.g., the number of estimated components for the mixing random variable A in the model X = AY). To our knowledge, no optimum detection algorithm has been obtained for such data. In an era when radiated ship noise is being dramatically reduced, a long-term investigation aimed at obtaining a likelihood ratio detector seems to be a very important task, provided that the data in this sample is representative of tactically or strategically important detection problems. Such investigations should be carried out by researchers familiar with the statistical theory and with access to a powerful computing facility. The problem is not apt to yield to either routine computational investigations or to purely theoretical research on stochastic processes. At the same time, regardless of the qualifications of investigators, the problem is too difficult to permit one to expect a rapid solution. Nevertheless, likelihood ratio detection for nonGaussian signals in Gaussian noise is an area of very little practical development. The NOSC Tape data indicates that it is an area that could be extremely important to the Navy. We strongly urge the Navy to consider a long-term research program on this problem.

I. STATISTICAL ANALYSIS OF NOSC TAPE

A. Randomness and Normality

The NOSC TAPE contained 15 min. of single-channel acoustical data (running from 11:40 to 11:55 am.) Sampling was done at 4096 hz, giving 3,686,400 data points for the 15 min. period. A sonargram of the data on the NOSC TAPE is presented in Figure 1.

Two 2 min. periods of time in the NOSC TAPE data were identified for analysis:

Period	Time	Feature
A	11:43:00-11:45:00	no target signal - noise only
B	11:47:30-11:49:30	target signal present

The data for these two periods were filtered and sampled and the results tested for normality and randomness as is now described.

The two sets of data (each comprised of 491,520 data points) for periods A and B were filtered using the following bandpass filters:

Filter	Frequency band
low	
mid	
high	

Thus, six sequences of 491,500 data points were produced; 3 for each time period. (A small number of data points were lost as part of filtering.) For each of these six sequences, every 40th data point was extracted to create six new sequences of length 12,288. Each of these sequences of data points was then segmented into 11 samples, each having 1000 data points. The final result was that, for each time period, 33 samples of length 1000 were formed; 11 from each filter for each of time periods A and B. This process is diagrammed in Figure 2 and Figure 3.

Figure 1

This page originally displayed a lofargram from the array which contained the hydrophone from which the data discussed here was obtained. This lofargram enabled us to determine regions where the target was definitely present, and regions where it was definitely absent. The signal data was primarily broadband in nature.

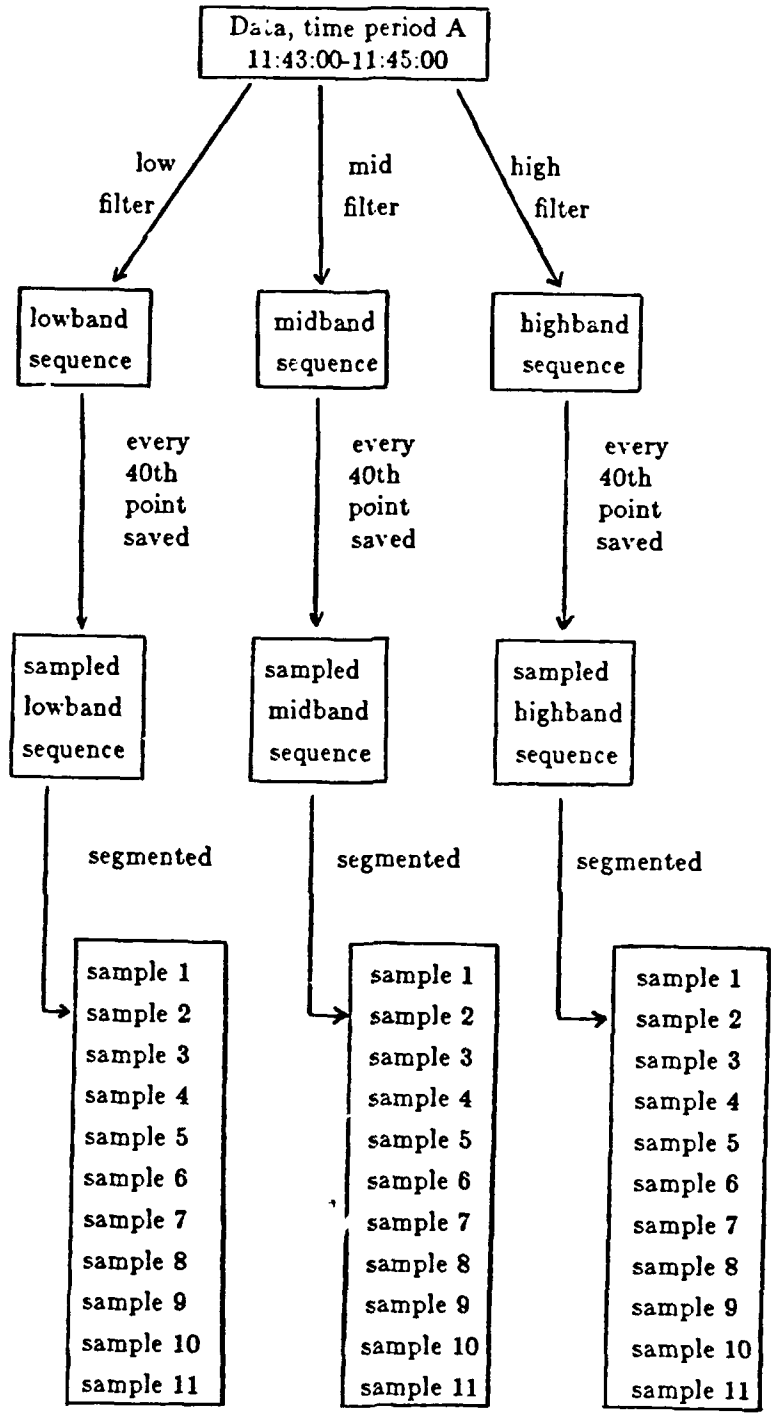


Figure 2. Processing of data in time period A. 33 samples of 1000 data points were obtained.

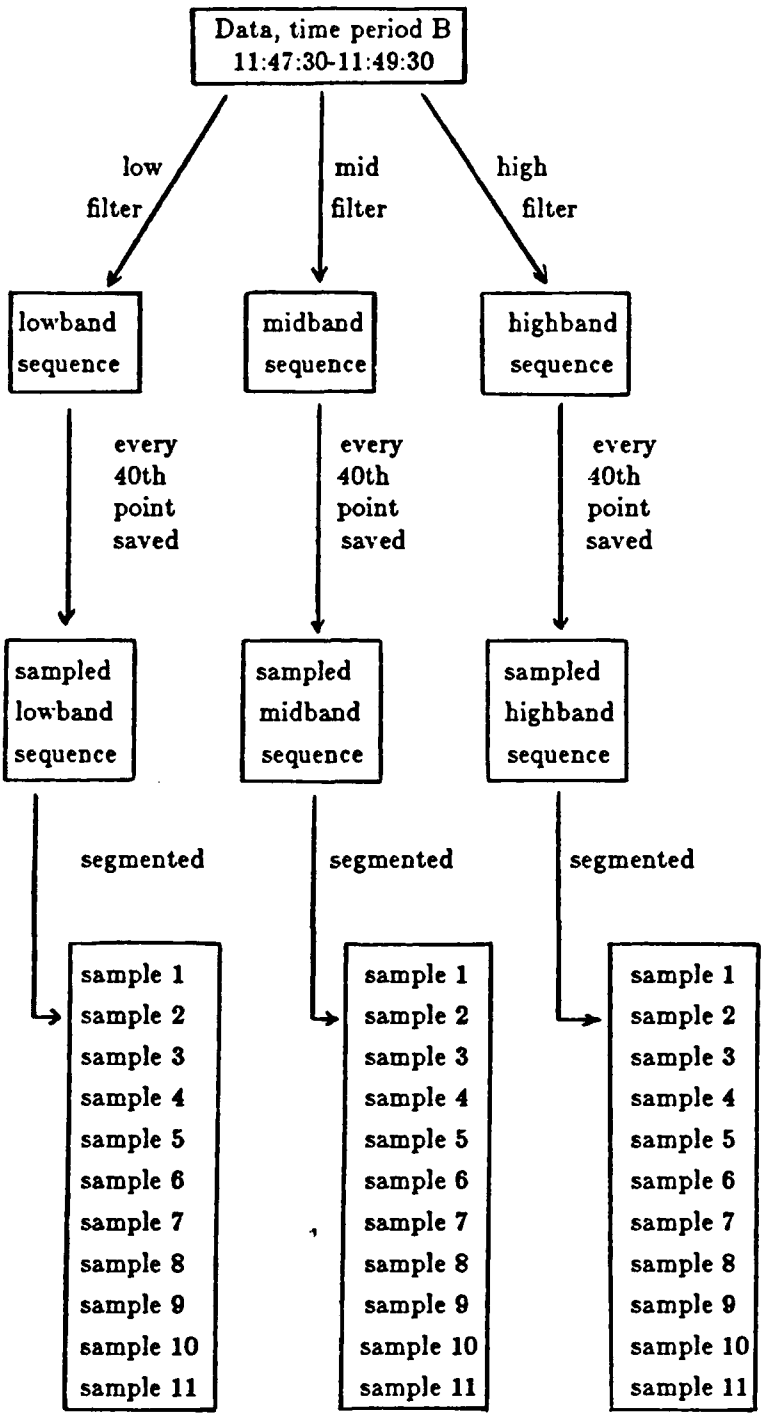


Figure 3. Processing of data in time period B. 33 samples of 1000 data points were obtained.

Four statistical tests of randomness were performed for each of the 66 samples. For each of the tests, the sample is reported to have failed (F) the test if the test p-value was less than 0.05. That is, the hypothesis of interest was rejected, with the probability of error not exceeding .05. Otherwise, the sample is reported to have passed (P). The four tests of randomness used were:

A - 'Runs-up-and-down' test
B - 'Runs-above-the-mean' test
C - 'Runs-above-the-mean' test (on absolute centered data)
D - Zero mean test

The results of these tests for each of the 66 samples can be found in Tables 1 and 2.

Four statistical tests of normality were performed for each of the 66 samples. For the first three tests, E, F, and G, the sample is reported to have failed (F) the test if the test p-value was less than 0.05. Otherwise, the sample is reported to have passed (P). For the fourth test, H, the sample is reported to have failed the test if the test p-value was less than 0.1. The four tests of normality used were:

E - Kolmogorov-Smirnov one-sample test
F - D'Agostino test
G - Chi-square test
H - Kurtosis test

The results of these tests for each of the 66 samples can be found in Tables 1 and 2. We define the nonnormality index for a sample to be the number of normality tests failed. Tables 1 and 2 also include the nonnormality index for each sample.

In cases where a sample failed two or more tests of normality, the sample data were posited to follow a mixture-of-normals (spherically-invariant) law. In these cases, the distribution of a mixture of normal random variables was fitted to the empirical distribution given by the sample. The algorithm used to make this fit is outlined in Appendix 1. The last column of Tables 1 and 2 gives the estimated number of normal components for each sample. Thus, if a sample failed less than two tests of normality then it was considered to be normal and the number of normal components entered in Tables 1 and 2 for the sample is 1. Where the nonnormality index was 2 or more, the number of normal components entered in Tables 1 and 2 (in the last column) is that found by the fitting algorithm. Table 3 presents summary statistics of the number of normal components found for each sample. A mixture of as many as seven independent normal random variables was sometimes needed to achieve an adequate fit. The normal density and the mixture-of-normals density found by the fitting algorithm are plotted against the empirical density for three different samples in Figures 4, 5, 6, and 7.

Samples		Tests of Randomness A B C D	Tests of Normality E F G H	Nonnormality Index	No. of Mixture Components
	1	passed all	passed all	0	1
	2	passed all	passed all	0	1
	3	passed all	P F P F	2	1
	4	passed all	passed all	0	1
Lowband	5	passed all	P F P F	2	4
samples	6	passed all	P F F F	3	1
	7	P P F P	passed all	0	1
	8	passed all	passed all	0	1
	9	passed all	F F F F	4	1
	10	passed all	P F P F	2	1
	11	passed all	passed all	0	1
	1	passed all	P F P F	2	4
	2	passed all	passed all	0	1
	3	passed all	passed all	0	1
	4	passed all	P F P F	2	3
Midband	5	F P P P	passed all	0	1
samples	6	passed all	P F P F	2	2
	7	P F P P	P P F P	1	1
	8	F P P P	passed all	0	1
	9	passed all	F F F F	4	1
	10	passed all	passed all	0	1
	11	passed all	passed all	0	1
	1	passed all	P F P F	2	4
	2	P P F P	passed all	0	1
	3	P F P P	P F F F	3	4
	4	passed all	P F P F	2	1
Highband	5	P P F P	passed all	0	1
samples	6	passed all	P F P F	2	2
	7	passed all	P F F F	3	1
	8	P P F P	passed all	0	1
	9	passed all	F F F F	4	2
	10	passed all	P F P F	2	1
	11	passed all	P F P F	2	5

Table 1. Results of analysis of NOSC Tape, time period A.

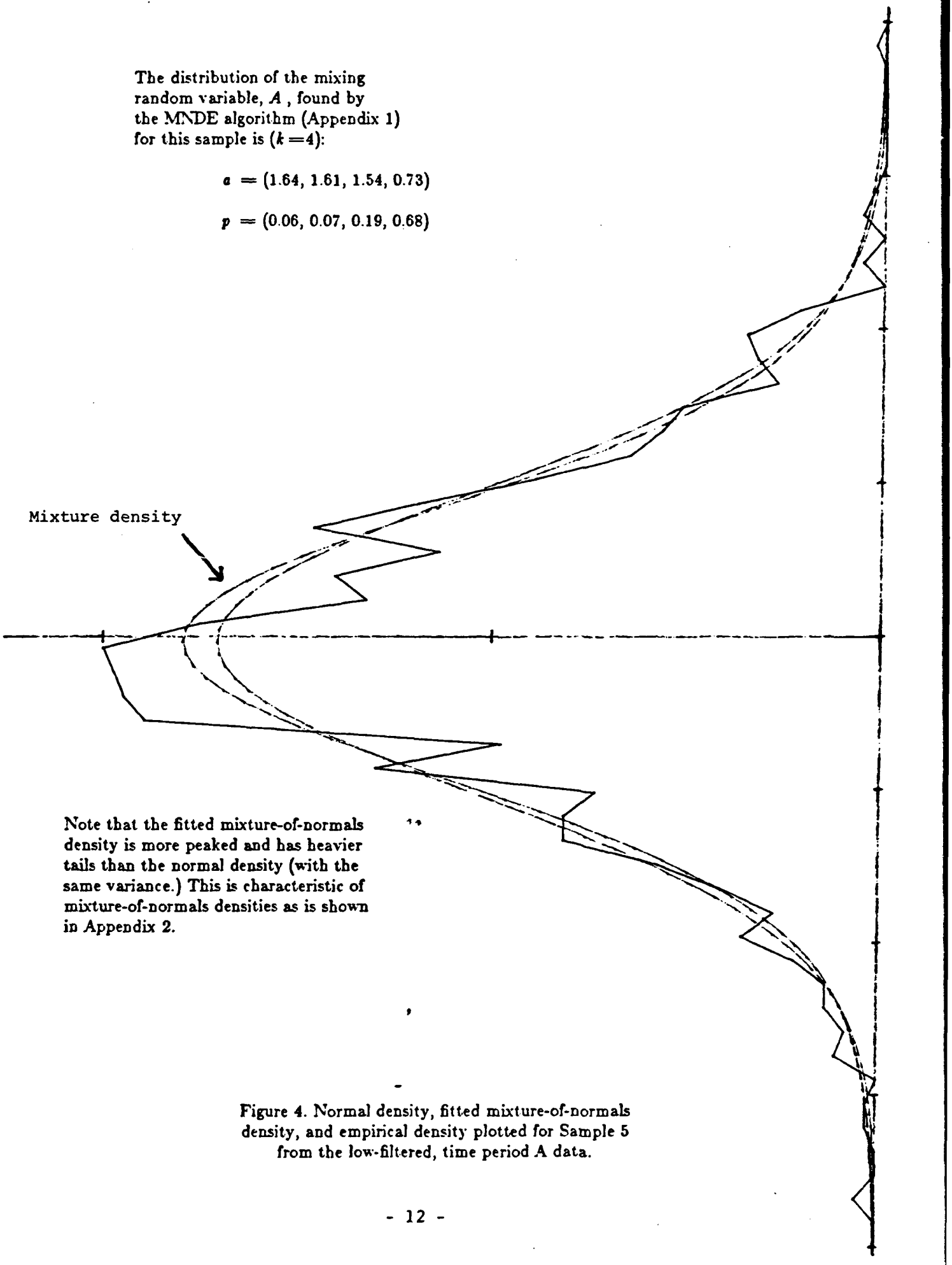
Samples		Tests of Randomness A B C D	Tests of Normality E F G H	Nonnormality Index	No. of Mixture Components
	1	F F P P	P F P F	2	7
	2	F F P P	P F P F	2	2
	3	F F F P	passed all	0	1
	4	F F P P	P F P F	2	6
Lowband	5	F F P P	passed all	0	1
samples	6	passed all	P F P F	2	4
	7	P P F P	P F P F	2	4
	8	F F F P	P F P F	2	6
	9	passed all	P F P F	2	2
	10	F F P P	P F P F	2	3
	11	F F P F	P F F F	3	2
	1	F F P P	P F P F	2	5
	2	F F P F	P F P F	2	2
	3	F P P P	passed all	0	1
	4	F F P P	P P P F	1	1
Midband	5	passed all	P F P F	2	1
samples	6	passed all	P F P F	2	2
	7	P F P P	P F P F	2	3
	8	passed all	P F P F	2	5
	9	F P P P	P F P F	2	2
	10	P F F P	P F P F	2	1
	11	F F P P	P F P F	2	2
	1	F P P P	P F P F	2	2
	2	passed all	P P P F	1	1
	3	passed all	passed all	0	1
	4	passed all	passed all	0	1
Highband	5	P P P F	passed all	0	1
samples	6	passed all	passed all	0	1
	7	passed all	passed all	0	1
	8	passed all	passed all	0	1
	9	P F P P	P F F F	3	5
	10	P P F P	passed all	0	1
	11	P F P P	P F F F	3	4

Table 2. Results of analysis of NOSC Tape, time period B.

The distribution of the mixing random variable, A, found by the MNDE algorithm (Appendix 1) for this sample is ($k = 4$):

$$a = (1.64, 1.61, 1.54, 0.73)$$

$$p = (0.06, 0.07, 0.19, 0.68)$$



Note that the fitted mixture-of-normals density is more peaked and has heavier tails than the normal density (with the same variance.) This is characteristic of mixture-of-normals densities as is shown in Appendix 2.

Figure 4. Normal density, fitted mixture-of-normals density, and empirical density plotted for Sample 5 from the low-filtered, time period A data.

The distribution of the mixing random variable, A, found by the MNDE algorithm (Appendix 1) for this sample is ($k = 5$):

$$a = (1.20, 1.22, 1.19, 1.15, 0.70)$$

$$p = (0.05, 0.10, 0.12, 0.24, 0.49)$$

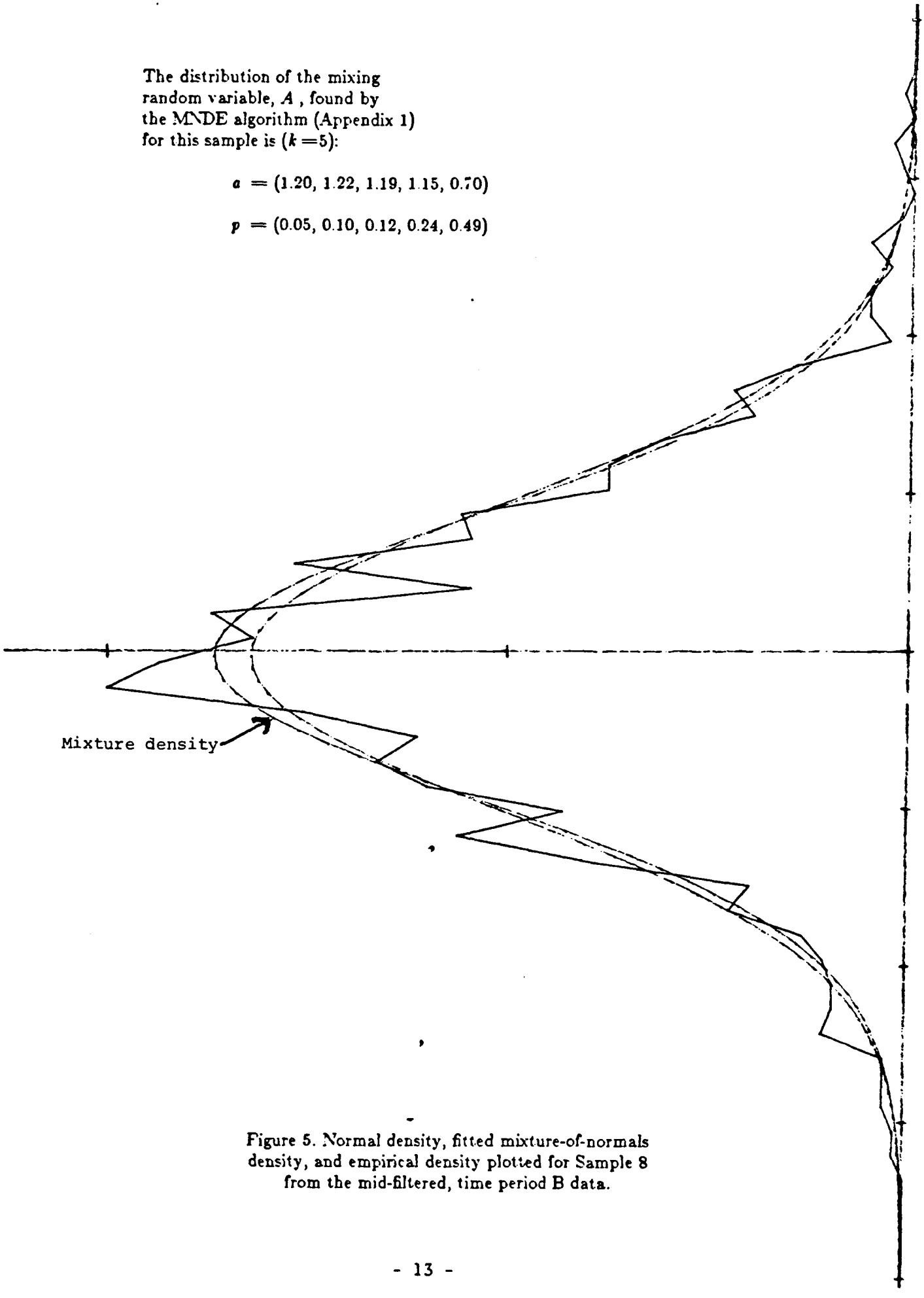


Figure 5. Normal density, fitted mixture-of-normals density, and empirical density plotted for Sample 8 from the mid-filtered, time period B data.

The distribution of the mixing random variable, A, found by the MNDE algorithm (Appendix 1) for this sample is ($k = 2$):

$$a = (1.59, 0.91)$$

$$p = (0.10, 0.90)$$

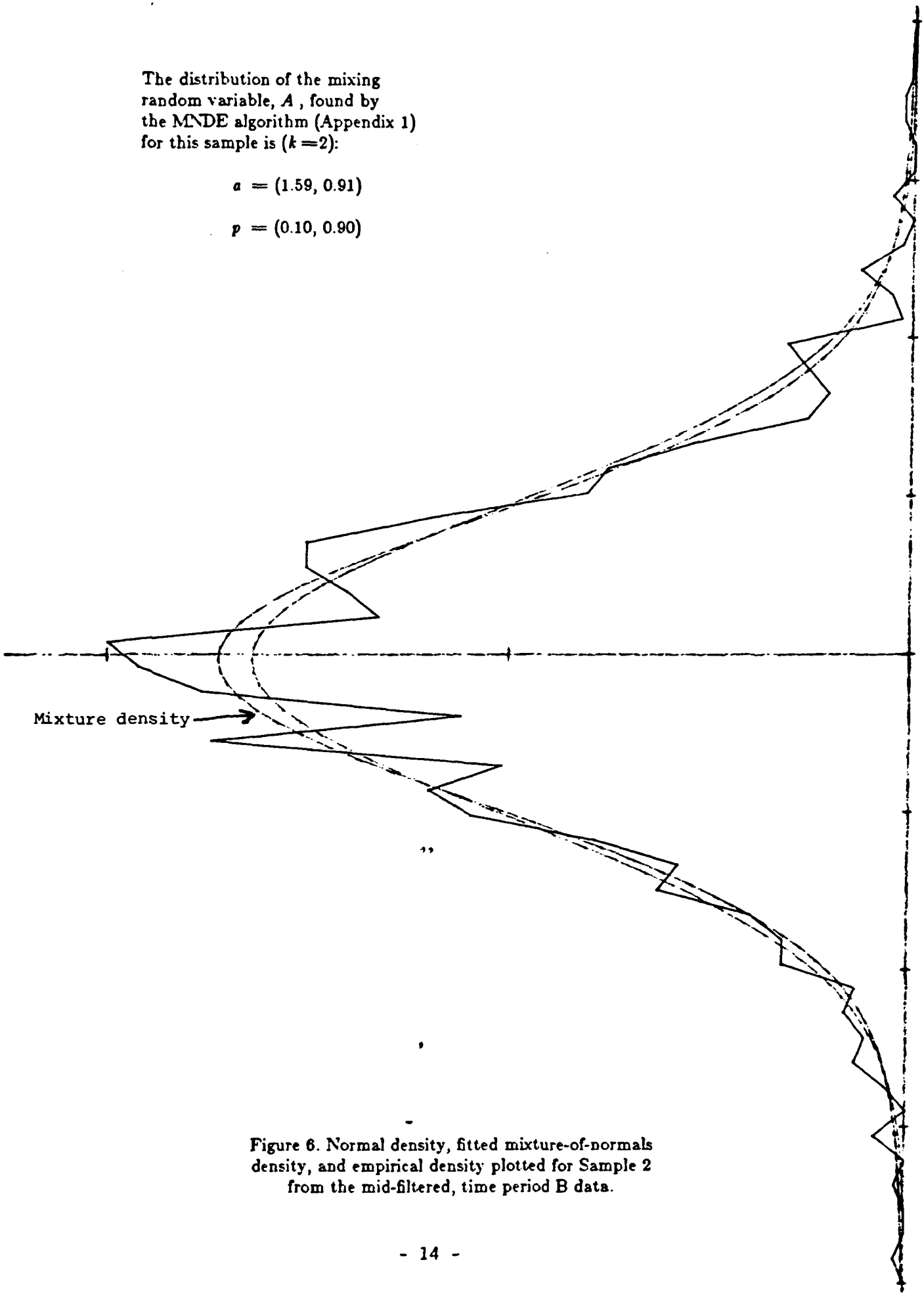


Figure 6. Normal density, fitted mixture-of-normals density, and empirical density plotted for Sample 2 from the mid-filtered, time period B data.

The distribution of the mixing random variable, A, found by the MNDE algorithm (Appendix 1) for this sample is ($k = 4$):

$$a = (3.30, 0.96, 0.93, 0.83)$$

$$p = (0.06, 0.08, 0.20, 0.66)$$

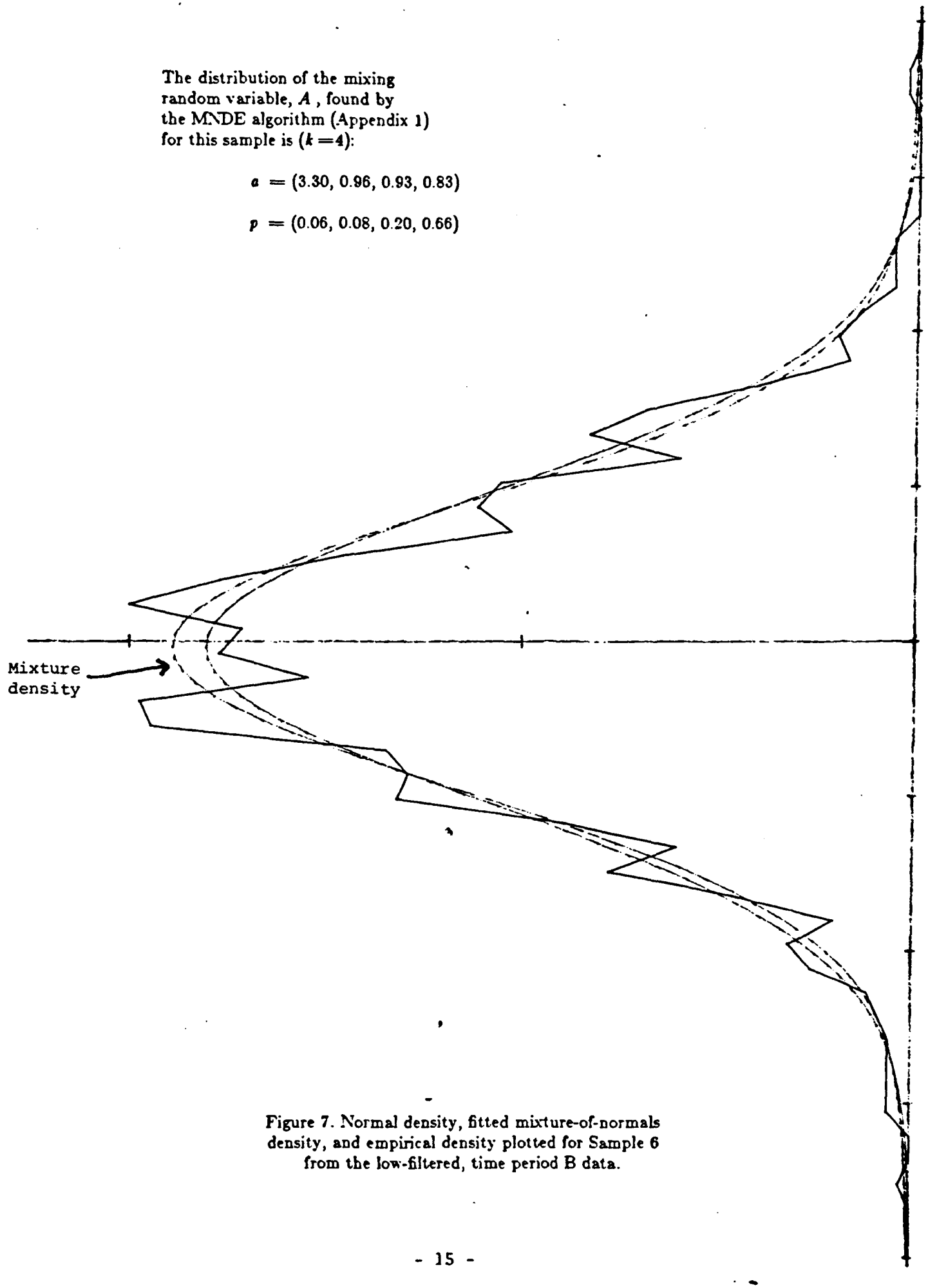


Figure 7. Normal density, fitted mixture-of-normals density, and empirical density plotted for Sample 6 from the low-filtered, time period B data.

	Time period A (noise only)	Time period B (target signal present)
Lowband	Mean=1.3 Median=1 Maximum=4	Mean=3.5 Median=3 Maximum=7
Midband	Mean=1.6 Median=1 Maximum=4	Mean=2.3 Median=2 Maximum=5
Highband	Mean=2.1 Median=1 Maximum=5	Mean=1.7 Median=1 Maximum=5

Table 3. Summary statistics of no. of normal components needed to adequately fit distributions of data samples.

The results of Tables 1 and 2 together with the summary statistics in Table 3 support the following conclusions:

The noise-only signal (time period A) is Gaussian in the low- and mid-frequency bands,

In the presence of a target (time period B), the received signal shows regular and significant departures from normality - particularly in the low- and mid-frequency bands,

Small, and apparently similar, departures from normality are present in both time periods in the high-frequency band,

II. DETECTOR PERFORMANCE USING NOSC TAPE

The sonargram corresponding to the 15 min. of data on the NOSC TAPE suggests the presence of a target signal during the period of time 11:47:00 - 11:52:00. Based upon this interpretation of the sonargram, the data were separated into three segments:

Segment	Time	Number of Data Points	Interpretation
1	11:40-11:47	1,720,320	noise only
2	11:47-11:52	1,228,800	signal present
3	11:52-11:55	737,280	noise only

Each segment of data was then divided in two by separating out every other data point. This resulted in six sets of data: (A small amount of data was discarded in this process to produce round numbers of data in each set.)

Set	Parent Segment	Number of Data Points	Interpretation
1	1	850,000	noise only
2	1	850,000	noise only
3	2	600,000	signal present
4	2	600,000	signal present
5	3	350,000	noise only
6	3	350,000	noise only

Next, sets 1 and 5 were joined together as were sets 2 and 6. The remaining four sets of data were then segmented into groups of 100 data points to form vectors of length 100. Thus the data were organized into the following final form:

(New) Set	Parent Sets	Number of Vectors	Interpretation
N1	1,5	12000	noise only
N2	2,6	12000	noise only
S1	3	6000	signal present
S2	4	6000	signal present

Data sets S2 and N2 were used to estimate the signal (plus noise) covariance, R_{S+N}, and noise covariance, R_N, matrices, respectively, while data sets S1 and N1 were reserved to directly test the performance of four signal detectors. In particular, data set S1 was used to calculate the mean signal (plus noise) vector, m_{S+N}, and data set N1 was used to calculate the mean noise vector,

m_N. See Table 4.

Estimates	Data set used
Mean signal vector, m_{S+N}	S1
Mean noise vector, m_N	N1
Signal covariance matrix, R_{S+N}	S2
Noise covariance matrix, R_N	N2

Table 4.

Using the estimates of noise and signal covariance and means, the signal-to-noise ratio (SNR) can be calculated through the following formula:

$$SNR = \frac{Trace\,(R_{S+N} + m_{S+N}\, m_{S+N}^T)}{Trace\,(R_N + m_N\, m_N^T)} - 1$$

The signal-to-noise ratio was found to be, SNR = 1.27.

For a given input vector x, a signal detector indicates "signal present" if and only if its output (the value of the detector test statistic) equals or exceeds a threshold; otherwise the detector decision is "noise only". The performance of four signal detectors was tested. These four signal detectors and their corresponding test statistics are presented in Table 5.

Signal Detector	Detector Test Statistic
Deflection	$x^T R_N^{-1} R_{S+N} R_N^{-1} x$
Matched filter	$m_{S+N}^T R_N^{-1} x$
Gauss vs. Gauss	$(x - m_N)^T R_N^{-1}(x - m_N)$ $- (x - m_{S+N})^T R_{S+N}^{-1}(x - m_{S+N})$
Constant False Alarm Probability (CFAP)	$\frac{m_{S+N}^T R_N^{-1} x}{\sqrt{x^T R_N^{-1} x}}$

Table 5. Detector test statistics.

Detection thresholds corresponding to selected false alarm probabilities are set based upon detector performance with signals known to contain only noise. Data set N1 was thus used to set detection thresholds for each of the four signal detectors.

Having found thresholds corresponding to selected false alarm probabilities, detector probability of detection can then be determined as the proportion of detector outputs which lead to the decision "signal present" when, indeed, a signal is present in the noise at the detector input. Data set S1 was used to find these detection probabilities.

Detector performance is represented by detection probability as a function of false alarm probability. Using data sets S1 and N1 as just described, detection probabilities were found for chosen false alarm probabilities for each of the four detectors. The results are plotted in Figure 8. The numerical probabilities from which the plots were constructed are tabulated in Appendix 3.

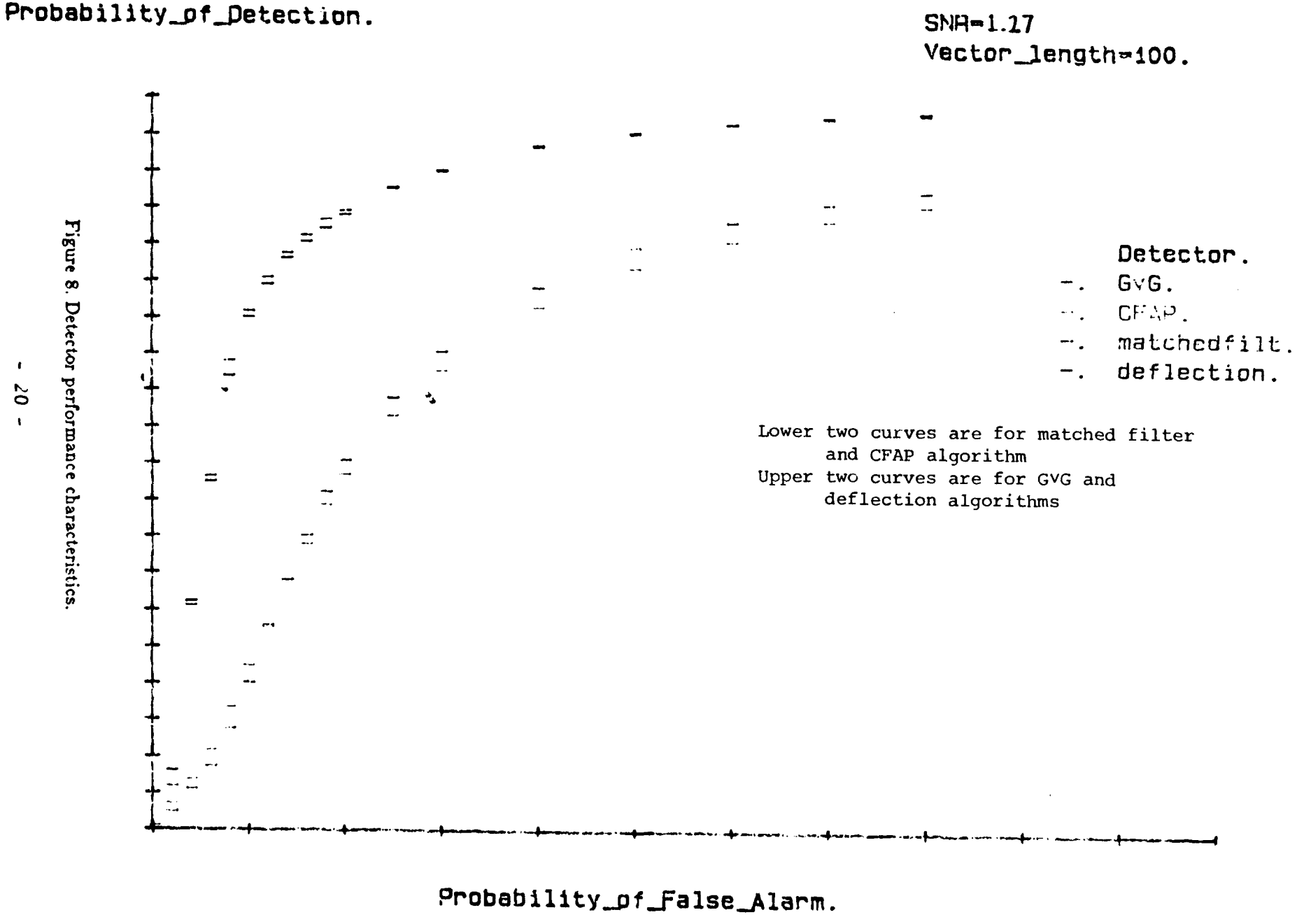


Figure 8. Detector performance characteristics.

Appendix 1

Mixture-of-Normals Distribution Estimation Algorithm

The mixture-of-normals distribution estimation (MNDE) algorithm described here fits a mixture-of-normals distribution to a random sample of size 1000 of standardized (subtract sample mean, divide by sample standard deviation) data.

The MNDE algorithm considers distributions of mixture-of-normals random variables of the form AX where A, X are independent random variables, X is N(0,1) and A is positive with finite support

$$\{a(1), a(2), ..., a(k)\}.$$

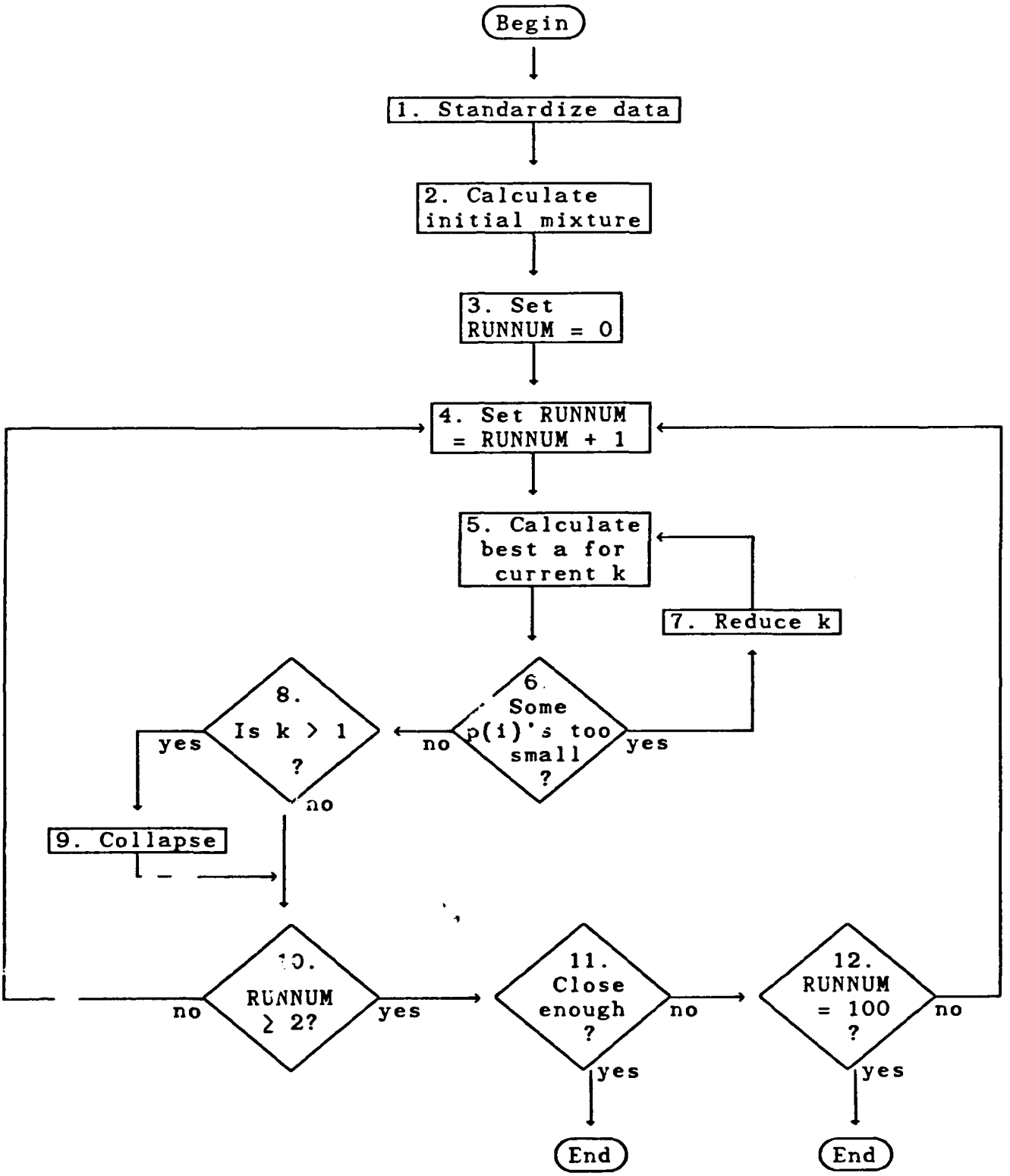
We write $p(i) = P\{A = a(i)\}$, $i = 1,..,k$, and define a, p to be the vectors

$$a = (a(1), a(2), ..., a(k)),$$
$$p = (p(1), p(2), ..., p(k)).$$

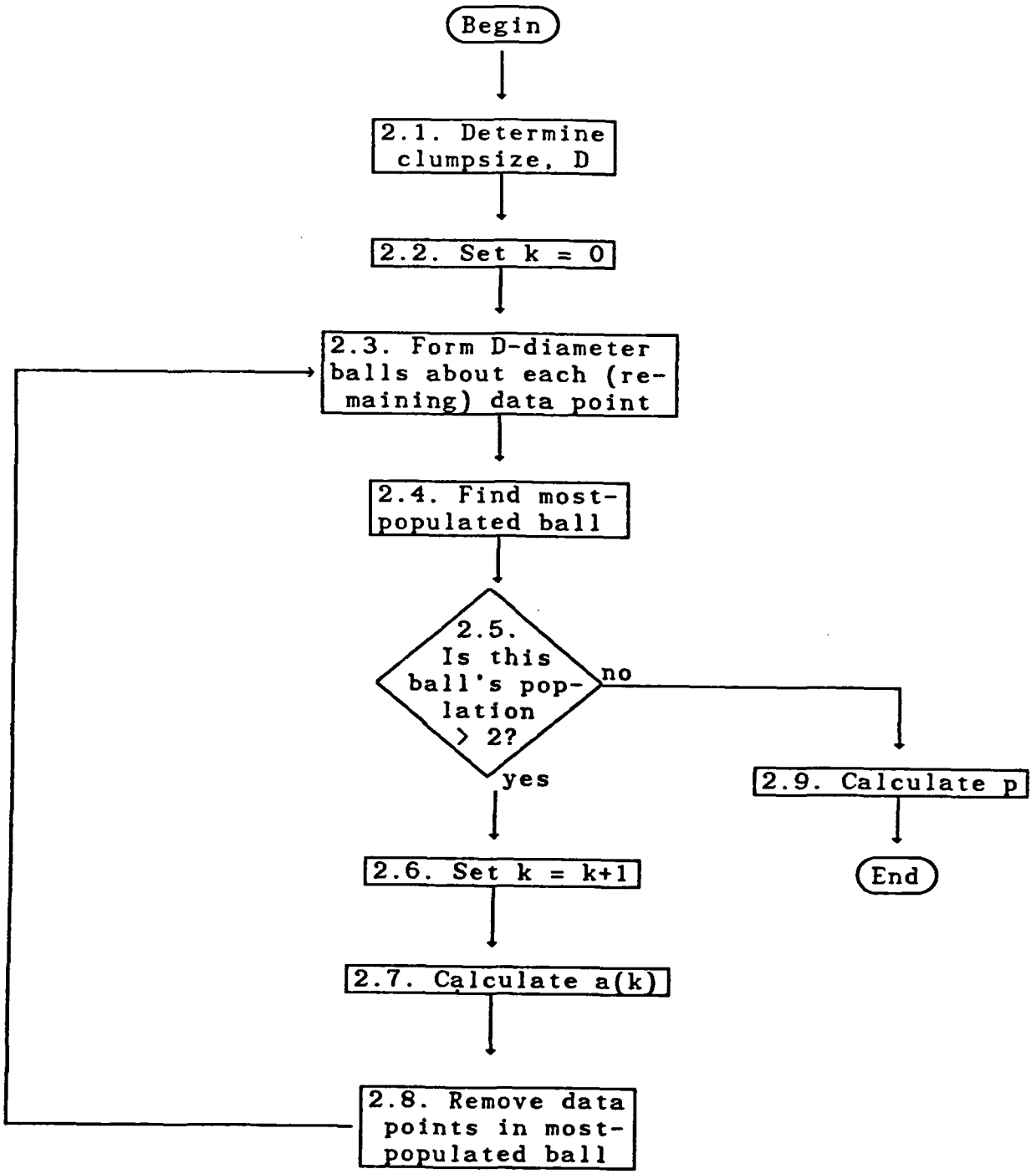
A is called a mixing random variable and the $p(i)$'s are called the mixing proportions of the distribution. Given an input, then, of 1000 data points, the MNDE algorithm outputs estimates of k, a, p.

The MNDE algorithm excludes from consideration mixing random variables, A, with mixing proportions $p(i) < 0.05$. Thus A can have no more than 20 mixing components, $a(i)$; i.e., $k > 20$ is not allowed.

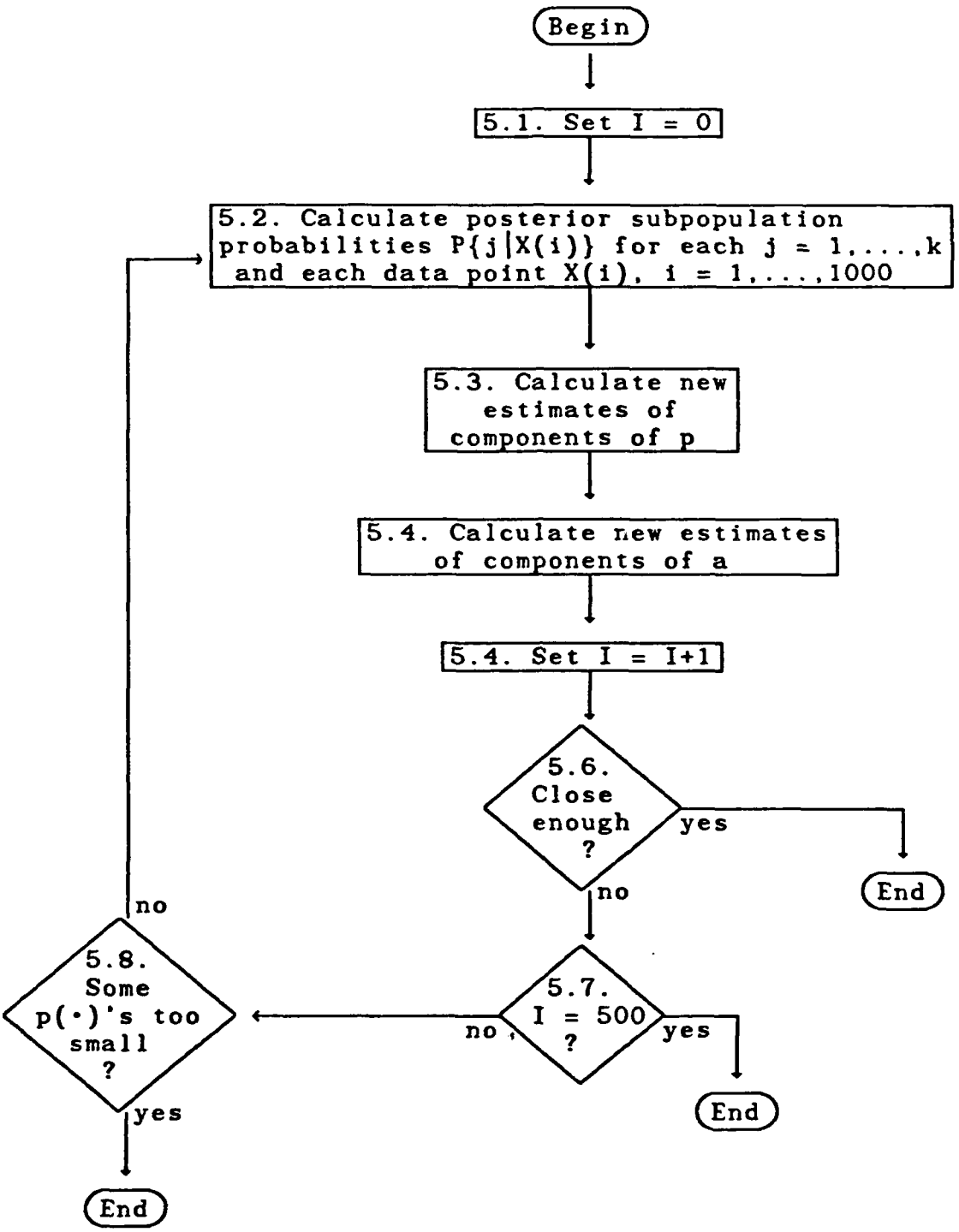
The MNDE algorithm is diagrammed in the following hierarchy of flowcharts. Following the flowcharts are remarks and explanations coded to the steps in the flowcharts. Note that succeeding flowcharts are coded to those preceeding.



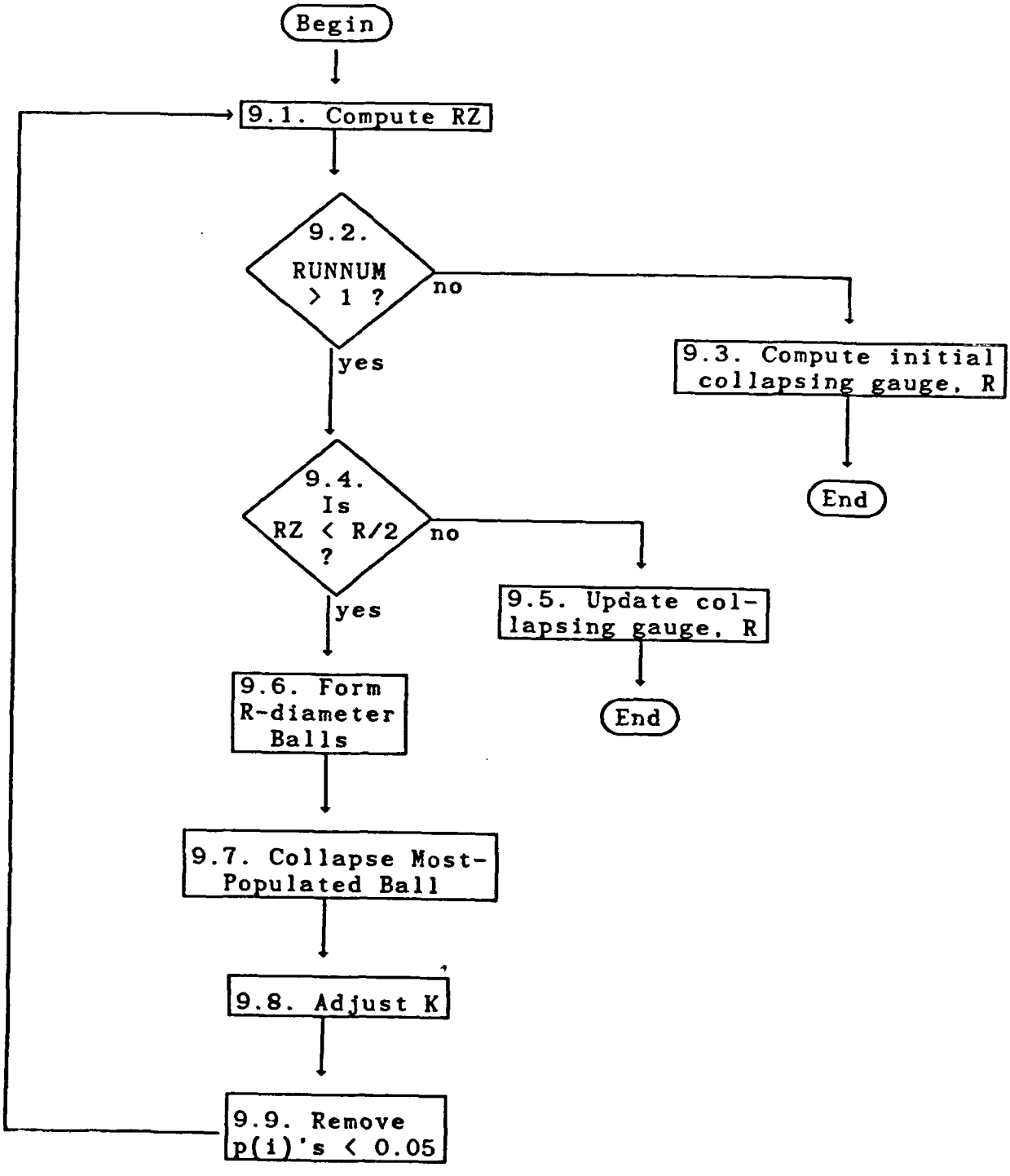
Flow Chart 1. MNDE Algorithm.



Flow Chart 2. Step 2. Calculate initial mixture.



Flow Chart 3. Step 5. Calculate best a for current k.



Flow Chart 4. Step 9. Collapse.

2. In Step 2 initial values are found for k and the vectors, a ,p .

2.1. Tracing through Step 2 of the algorithm reveals that the initial value of k is larger for smaller clumpsizes, D .

2.7. $a(k)$ is calculated as the average of the data points in the current most-populated ball.

2.9. The components $p(i)$, $i=1,...,k$, of the vector p are found by

$$p(i) = \frac{n(i)}{N}$$

where

$$N = n(1)+n(2)+...+n(k)$$

and where $n(i)$, $i=1,...,k$, are the populations of the k most-populated balls as found in Step 2.4.

5.2. The conditional probabilities $P\{j \mid X(i)\}$ that the data point $X(i)$ comes from the subpopulation characterized by $A = a(j)$ are calculated, for each $j=1,...,k$, and each $X(i)$, $i=1,...,1000$, by

$$P\{j \mid X(i)\} = \frac{p(j)f(X(i) \mid A=a(j))}{f(X(i))}$$

where $f(\ \mid A=a(j))$ is the normal density, $N(0,a^2(j))$, and $f(\ )$ is the density of AX.

5.3. New estimates of the components $p(j)$, $j=1,...,k$, are calculated by

$$p(j) = \frac{1}{1000}\sum_{i=0}^{1000} P\{j \mid X(i)\}.$$

5.4. New estimates of the components $a(j)$, $j=1,...,k$, are calculated by

$$a^2(j) = \frac{1}{1000p(j)}\sum_{i=1}^{1000} \frac{X^2(i)}{\sigma^2} P\{j \mid X(i)\}.$$

5.6. Two successive estimates a , p and a' , p' (k is not changing here) are considered close if

$$\sum_{i=1}^{k}[p'(i)-p(i)]^2 + \sum_{i=1}^{k}\left[\frac{a'(i)}{s'} - \frac{a(i)}{s}\right] < 2\times10^{-6}k$$

where

$$s = \sum_{i=1}^{k} a(i),$$

$$s' = \sum_{i=1}^{k} a'(i).$$

5.7. Step 5 has never been found to terminate as a result of the condition: $I = 500$.

5.8. If any component of p is less than 0.0001 then a flag is set. This flag is later checked in Step 6.

6. A mixing proportion, $p(i)$, is too small if $p(i) < 0.0001$.

9.1. RZ is the minimum nearest-neighbor distance of $a(i)$, $i=1,...,k$.

9.3. The collapsing gauge, R, is used to decide whether or not certain of the $a(i)$'s are close enough to be merged together. Initially, if $RZ > 0.0002$, we set $R = RZ$; otherwise we set R=0.005. This ensures $R \geq 0.0002$ and prevents an initially small RZ from blocking collapsing in following iterations.

9.5. We set $R = \max\{R, RZ\}$.

11. The results, k, a, p, of the current iteration and the results, k', a', p' of the previous iteration are "close" if both of the following inequalities are satisfied:

$$\left| \frac{k' - k}{k} \right| \leq 0.1,$$

$$\sum_{j=1}^{k} (p(j) - p'(j))^2 + \sum_{j=1}^{k} \left[\frac{a(j)}{|a|} - \frac{a'(j)}{|a'|} \right]^2 \leq 0.0001k$$

where

$$|a| = \left[\sum_{j=1}^{k} a^2(j) \right]^{\frac{1}{2}},$$

$$|a'| = \left[\sum_{j=1}^{k} a'(j)^2 \right]^{\frac{1}{2}}.$$

12. For our samples, in running the MNDE algorithm, the number of iterations, *RUNNUM*, never exceeded 10.

Appendix 2

Characteristics of the Mixture-of-Normals Density

A mixture-of-normals random variable, $Y = AX$, where A, X are independent random variables, X is $N(0,\sigma^2)$, and A is positive with finite support, $\{a(1),..., a(k)\}$, has density

$$f_Y(y) = \sum_{i=1}^{k} \frac{p(i)}{\sigma a(i)\sqrt{2\pi}} \exp\left[-\frac{1}{2\sigma^2}\frac{y^2}{a^2(i)}\right]$$

Then

$$f_Y(0) = \frac{1}{\sigma\sqrt{2\pi}} E\left[\frac{1}{A}\right]$$

and, using Jensen's inequality followed by Schwarz's inequality,

$$E\left[\frac{1}{A}\right] > \frac{1}{E[A]} > \frac{1}{\sqrt{E[A^2]}}.$$

(The inequalities are strict provided A is nondegenerate.) Now $E[Y^2] = E[A^2X^2] = E[A^2]\sigma^2$ so that $E[A^2] = 1$ for $Var[Y] = \sigma^2$. Then

$$f_Y(0) > \frac{1}{\sigma\sqrt{2\pi}}.$$

Thus the density of a mixture-of-normals random variable, $Y = AX$, is more peaked than that of a normal random variable with the same variance.

To see that the density of $Y = AX$ has heavier tails than those of the density of a normal random variable with the same variance, consider the limit

$$\lim_{x\to\infty} \frac{f_Y(x)}{f(x)}$$

where $f(x)$ is the $N(0,\sigma^2)$ density,

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}.$$

We have

$$\lim_{x\to\infty}\frac{f_Y(x)}{f(x)}=\lim_{x\to\infty}\sum_{i=1}^{k}\frac{p(i)}{a(i)}\exp\left[-\frac{1}{2\sigma^2}\frac{x^2}{a^2(i)}+\frac{x^2}{2\sigma^2}\right]$$

For at least one term in the above sum, $a(i)>1$, otherwise $E[A^2]\neq 1$ unless $A=1$ with probability 1. So,

$$\lim_{x\to\infty}\frac{f_Y(x)}{f(x)}=\infty$$

and we conclude that, for nondegenerate mixing random variables, A, $Y=AX$ has heavier tails than the corresponding normal random variable (having the same variance.)

Appendix 3

Probability of Detection Data

The data in Table 6 were used to construct the curves presented in Figure 8. The probabilities of detection given in Table 6 for each of the four signal detectors were determined by the proportion of times signals from data set S1 caused the signal detector output to exceed the detection threshold. The detection thresholds were, in turn, set using noise signals from data set N1 to maintain the false alarm probability at the prescribed value.

False Alarm Probability	Probability of Detection			
	CFAP	Matched Filter	Deflection	Gauss vs. Gauss
0.001	0.0055	0.0008	0.0002	0.0002
0.010	0.0358	0.0244	0.0817	0.0590
0.020	0.0692	0.0553	0.3117	0.3055
0.030	0.1087	0.0887	0.4758	0.4822
0.040	0.1673	0.1371	0.6192	0.6397
0.050	0.2251	0.2014	0.7008	0.7080
0.060	0.2803	0.2775	0.7453	0.7538
0.070	0.3405	0.3403	0.7808	0.7867
0.080	0.3898	0.4014	0.8037	0.8115
0.090	0.4442	0.4617	0.8220	0.8332
0.100	0.4863	0.5049	0.8410	0.8460
0.125	0.5658	0.5894	0.8778	0.8785
0.150	0.6283	0.6542	0.9013	0.9018
0.200	0.7158	0.7421	0.9367	0.9340
0.250	0.7689	0.7970	0.9542	0.9517
0.300	0.8064	0.8326	0.9673	0.9662
0.350	0.8352	0.8578	0.9765	0.9742
0.400	0.8561	0.8752	0.9832	0.9798

Table 6. Detector performance data.

References

[1] C.R. Baker and A.F. Gualtierotti, Likelihood ratios and signal detection for nonGaussian processes, *Stochastic Processes in Underwater Acoustics,* Lecture Notes in Control and Information Science, Springer-Verlag, 85, 154-180, Berlin (1986).

[2] D. Middleton, Statistical-physical models of electromagnetic interference, *IEEE Trans. on Electromagn. Compat., EMC-19,* 106-127 (1977).

[3] C.R. Baker, Optimum quadratic detection of a random vector in Gaussian noise, *IEEE Trans. on Communications Technology, 14,* 802-805 (1966).

[4] C.R. Baker and A.F. Gualtierotti, Discrimination with respect to a Gaussian process, *Probability Theory and Related Fields, 71,* 159-182 (1986).